



Published in final edited form as:

Obesity (Silver Spring). 2016 October ; 24(10): 2057–2063. doi:10.1002/oby.21603.

Reproducibility assessment of brain responses to visual food stimuli in adults with overweight and obesity

R Drew Sayer^a, Gregory G Tamer Jr^b, Ningning Chen^c, Jason R Tregellas^d, Marc-Andre Cornier^e, David A Kareken^f, Thomas M Talavage^b, Megan A McCrory^g, and Wayne W Campbell^a

^aDepartment of Nutrition Science, Purdue University, West Lafayette, IN

^bWeldon School of Biomedical Engineering, Purdue University, West Lafayette, IN

^cDepartment of Statistics, Purdue University, West Lafayette, IN

^dDepartment of Psychiatry, University of Colorado School of Medicine, Anschutz Medical Campus, Aurora, CO

^eDepartment of Medicine, Division of Endocrinology, Metabolism and Diabetes, University of Colorado School of Medicine, Anschutz Medical Campus, Aurora, CO

^fDepartment of Neurology, Indiana University School of Medicine, Indianapolis, IN

^gDepartment of Nutrition, Georgia State University, Atlanta GA

Abstract

Objective—The brain's reward system influences ingestive behavior and subsequently, obesity risk. Functional magnetic resonance imaging (fMRI) is a common method for investigating brain reward function. We sought to assess the reproducibility of fasting-state brain responses to visual food stimuli using BOLD fMRI.

Methods—*A priori* brain regions of interest included bilateral insula, amygdala, orbitofrontal cortex, caudate, and putamen. Fasting-state fMRI and appetite assessments were completed by 28 women (n=16) and men (n=12) with overweight or obesity on 2 days. Reproducibility was assessed by comparing mean fasting-state brain responses and measuring test-retest reliability of these responses on the 2 testing days.

Results—Mean fasting-state brain responses on Day 2 were reduced compared to Day 1 in the left insula and right amygdala, but mean Day 1 and Day 2 responses were not different in the other regions of interest. With the exception of the left orbitofrontal cortex response (fair reliability), test-retest reliabilities of brain responses were poor or unreliable.

Corresponding Author: Wayne W Campbell; Purdue University, 700 West State Street, West Lafayette, IN 47907; Phone: 765-494-8236; Fax: 765-494-0674, campbellw@purdue.edu.

Disclosures: No conflicts of interest to declare.

Author Contributions: WWC, TMT, MAM, GGT, and RDS conceived the research project; DAK, MAC, and JRT assisted with further development of the study design; RDS was responsible for subject recruitment; RDS and GGT conducted the research; RDS, GGT, and NC performed the statistical analyses; RDS and WWC wrote the manuscript and GGT, NC, JRT, MAC, DAK, TMT, and MAM provided critical feedback and edits to the manuscript. All authors take responsibility for the final content of the manuscript.

Conclusion—fMRI-measured responses to visual food cues in adults with overweight or obesity show relatively good mean-level reproducibility, but considerable within-subject variability. Poor test-retest reliability reduces the likelihood of observing true correlations and increases the necessary sample sizes for studies.

Keywords

fMRI; reward; test-retest reliability; obesity; appetite

Introduction

The brain's reward system is an important modulator of ingestive behavior that subsequently influences obesity risk (1). Functional magnetic resonance imaging (fMRI) is commonly used to compare brain reward responses elicited by images of palatable and/or energy dense foods compared to nonfood images such as landscapes, animals, and/or household items (2–4). Meta-analytical findings from these studies demonstrate that images of food elicit robust responses in reward-associated brain regions – such as the insula, amygdala, orbitofrontal cortex, and striatum – (2) that are greater in adults with obesity relative to those with normal body weights (3,4).

These data have prompted investigators to consider potential interventions designed to modulate brain reward responses to food cues and, subsequently, ingestive behavior and obesity risk. For example, intervention studies have investigated potential modulation of food cue-induced responses by aerobic exercise (5–7), dietary protein (8,9), as well as administration of exogenous “appetite-related” hormones (10–12). However, data regarding the reproducibility of brain reward responses to visual food cues are lacking. These data are critical for the proper implementation and interpretation of intervention studies. The test-retest reliability of a measurement has implications relating to sample size determination and the maximum observable correlations among outcomes (13).

Therefore, the primary aim of this study was to assess the reproducibility of fMRI-measured responses to visual food cues in the insula, amygdala, orbitofrontal cortex, and dorsal striatum (caudate and putamen) in adults with overweight or obesity. These reward-related brain regions of interest were chosen based on previous work indicating their responsivity to visual food stimuli and greater observed responses in people with obesity (4). We hypothesized that 1) mean fasting brain responses in these reward-associated brain regions would not differ on 2 testing days with similar experimental conditions and 2) fasting-state neural responses would demonstrate good or excellent test-retest reliability (Intraclass Correlation Coefficient ≥ 0.60). The influences of time of day (morning vs. evening) and sex on fasting-state brain responses and appetite ratings were considered as secondary outcomes.

Methods

Subjects

Thirty-six individuals (20 female, 16 male) were recruited from the greater Lafayette, IN, community to participate in 1 of 3 fMRI research studies at Purdue University (Figure 1).

Thirty individuals (17 female, 13 male) completed fasting-state fMRI and appetite assessments on 2 testing days. Postprandial fMRI and appetite assessments were collected in the aforementioned trials, but those results are not presented here. Data from 1 male subject were excluded due to excessive head movement (> 2.5 mm) during fMRI scanning, and data from one female subject were excluded because she completed testing days at a different time of day (1 in the morning and 1 in the evening). The final sample included 28 individuals (16 female, 12 male) (Figure 1). Inclusion criteria for this study were: male or female; age 18 – 45 y; overweight or obese BMI (25.0 – 40.0 kg/m²); weight stable (± 3 kg for previous 6 mo); no tobacco use; no diabetes; not pregnant or lactating; not claustrophobic; no implanted pacemakers/automated defibrillators or ferromagnetic metal. All subjects provided written informed consent and received a monetary stipend. The consent form and all study procedures and documents were approved for use by the Purdue University Biomedical Institutional Review Board.

Experimental Design

Sixteen subjects completed 2 testing days in the morning (a.m.) beginning between 0700 and 0900 after a 10-hour overnight fast (no food or energy- or caffeine-containing beverages after 2100, 2200, or 2300 depending on the subject's scheduled start time). Twelve subjects completed 2 testing days in the evening (p.m.), which began at 1700, 5 hours after consuming a provided lunch (30% of estimated daily energy requirement (14)). Subjects completing p.m. testing days were asked to refrain from consuming any food or energy- or caffeine-containing beverages in the time between finishing lunch and their arrival at the MRI facility. For descriptive purposes, assessments completed before meal consumption on a.m. and p.m. testing days are referred to as "fasting-state." All testing days were separated by at least 3 days (mean: 18 days, range: 3 – 35 days) during which time the subjects consumed self-chosen food and beverages *ad libitum*.

Body Mass Index

Body mass was measured using a digital platform scale (Ohaus, ES200L, Toledo, OH, USA) and height was measured using a wall-mounted stadiometer (Holtain Ltd., Crymych, Wales, UK). Body mass index (kg/m²) was calculated using these measurements.

Appetite Assessments

Subjects rated their hunger, desire to eat, and fullness on continuous visual analog scales (VAS) (15) using Adaptive Visual Analog Scales software (16).

fMRI Data Acquisition

Functional imaging was performed using a 3.0 Tesla magnetic resonance scanner (General Electric, Signa HDx, Milwaukee, WI, USA) while subjects were lying quietly in a supine position and presented with visual stimuli using NordicNeuroLab's VisualSystem (Bergen, Norway). Visual stimuli consisted of images of food of high hedonic value and neutral nonfood-related objects (e.g. tools, landscapes, and household items), which were previously validated (17) and utilized in previously published reports (18–23). Head movement was limited by placing foam pads behind the subjects' necks and between the 16-channel head

coil (Nova Medical, Inc., Model NMSC-025A, Wilmington, MA) and all sides of the subjects' heads. A localizer scan was prescribed and centered at the subjects' brow line. The type, number and placement of foam pads, the location of subjects inside of the fMRI scanner, and the location of localizer prescription were noted on the first testing day and replicated to the greatest extent possible on the second testing day and for postprandial fMRI sessions.

Three functional runs were performed during each fMRI session. Each run consisted of 3 blocks of visual food stimuli and 3 blocks of visual nonfood stimuli presented in a pseudorandomized order using PsychoPy, Version 1.76.00 (24). Each block of visual stimuli lasted 30 seconds and included 10 images presented for 2.5 seconds each with a 0.5 second fade between each image. Blocks of a low-level baseline stimulus (fixation cross) lasting 16 seconds each were presented before the first visual stimuli block, in between blocks of visual stimuli, and after the final visual stimuli block.

Functional images were acquired with an echo-planar gradient-echo T2* blood oxygenation level dependent (BOLD) contrast sequence, with TR = 2000 ms, TE = 30 ms, 64² matrix, 20 cm² field of view, 40 slices to cover the whole brain, 3.1 mm thick, and no gap between slices. A high resolution, T1-weighted anatomical scan was completed after functional imaging for coregistration with functional images.

fMRI Data Processing and Analysis

fMRI data preprocessing and first-level fMRI data analyses (food vs nonfood BOLD contrasts) were completed using AFNI (<http://afni.nimh.nih.gov>) (25). The first five volumes of each functional run (presented during the fixation cross block) were excluded to eliminate any T1 relaxation effects that may have been present due the relatively short TR of 2000 ms. Functional runs were then slice-time corrected, motion corrected to the first non-excluded image of the first functional run, smoothed using a 4.0 mm Gaussian blur, and the signal was normalized. Motion and slice-time corrected functional runs were then aligned with the high resolution anatomical scan.

Censor files were created to identify volumes within each functional run with excessive head motion (>2.5 mm). Those volumes that were censored were excluded and the six (roll, pitch, yaw, left-right, superior-inferior, anterior-posterior) motion time-series were included as covariates in the first-level, food vs. nonfood contrast regression model. The resulting food vs. nonfood contrast was expressed as a β coefficient for each fMRI session.

We first examined *a priori* regions of interest (bilateral insula, amygdala, orbitofrontal cortex, caudate, and putamen) to search for local maxima from the food vs. nonfood BOLD contrast during the fasting fMRI session on the first testing day. Spherical regions of interest with a 3mm radii centered at the local maxima were then created to form each functional region of interest. Functional regions of interest obtained from the fasting fMRI scan on the first testing day were used again on the second testing day and for all postprandial fMRI scans.

Statistical Analysis

Mean β coefficients (average of all voxels in each region of interest) representing the first-level, food vs. nonfood BOLD contrasts were analyzed using single-sample Student's *t*-tests (SAS, Version 9.3, PROC TTEST) to determine if the contrasts were significantly different from zero (indicating a greater response to visual food vs. nonfood stimuli). A Bonferroni correction was applied to the results of the *t*-tests to correct for multiple comparisons among 10 *a priori* brain regions of interest ($\alpha = 0.05 / 10 = 0.005$).

Repeated measures ANOVA (SAS, Version 9.3, PROC MIXED) was used to assess the effects of testing day (Day 1 vs. Day 2, repeated factor), time of day (a.m. vs. p.m.), and sex on fasting brain reward responses (food vs. nonfood contrast) and appetite (hunger, desire to eat, and fullness) ratings. A Tukey-Kramer adjustment for multiple comparisons was utilized for the ANOVA tests. Test-retest reliabilities of fasting brain reward responses and appetite ratings were determined by 2-way, mixed effects model intraclass correlation coefficients [ICC(3,1)] (26) and Lin's correspondence correlation coefficients (CCC) (27) using IBM SPSS Statistics, Version 22. ICC(3,1) and CCC were interpreted as: < 0.40: poor reliability, 0.40 – 0.59: fair reliability, 0.60 – 0.74: good reliability, 0.75: excellent reliability (28). Negative ICC(3,1)s and CCCs were interpreted as being equivalent to zero and to represent complete unreliability (29).

Correlations among brain reward responses, body mass index, and appetite ratings were assessed using Pearson's correlation coefficients (SAS, Version 9.3, PROC CORR). Data are presented as mean \pm SEM and significance was set at $\alpha = 0.05$ unless otherwise noted.

Results

Subject Characteristics

On average, subjects were 27 years old and had a body mass index near the cutoff for obesity (Table 1).

fMRI Responses

Visual food stimuli presented in the fasting-state on Day 1 elicited greater responses compared to neutral nonfood stimuli ($P < 0.005$) in 9 of 10 *a priori* regions of interest: left and right insula, amygdala, orbitofrontal cortex, and caudate, and right putamen (Figure 2, Table 2). On Day 2, greater responses to visual food vs. nonfood stimuli were observed in all *a priori* brain regions of interest except the right caudate ($4.58 \times 10^{-4} \pm 1.32 \times 10^{-4}$, $P = 0.0071$) (Table 2). According to the ANOVA model, fasting-state responses to visual food stimuli were greater on Day 1 compared to Day 2 in the right amygdala and there was a trend for left insula responses to be greater on Day 1 ($P = 0.076$) (Table 3). No responses were greater on Day 2 vs. Day 1. A voxel-wise subtraction analysis also suggested differences in Day 1 vs. Day 2 responses in the left insula, but differences in right amygdala responses were not further substantiated by this secondary analysis (Figure 3). Fasting-state responses were not different on Days 1 and 2 in the remaining regions of interest (Table 3). Although the left insula and right amygdala responses were attenuated on Day 2, the food vs. nonfood contrasts remained significant in both regions (left insula β : $9.37 \times 10^{-4} \pm 1.50$

$\times 10^{-4}$, $P < 0.0001$; right amygdala β : $5.78 \times 10^{-4} \pm 1.82 \times 10^{-4}$, $P = 0.0025$). There was a trend for a greater response in the right amygdala of females compared to males ($P = 0.053$). Similarly, there were trends for greater responses in a.m. compared to p.m. sessions in the right insula ($P = 0.065$) and right putamen ($P = 0.077$).

The left orbitofrontal response to visual food stimuli demonstrated fair test-retest reliability. Reliabilities of responses in the other regions of interest were poor or unreliable, although left insula and left amygdala responses were near the cutoff for fair reliability (ICC(3,1): 0.389 and 0.390, respectively) (Table 4). Visual representations of differences in Day 1 and Day 2 brain responses in *a priori* brain regions of interest for each subject are available as supporting material online (Figures S1 – S10).

Appetite Ratings

Mean fasting-state appetite ratings were not different on Day 1 vs. Day 2, respectively (hunger: 50 ± 5 vs. 49 ± 4 mm; desire to eat: 51 ± 4 vs. 52 ± 4 mm; fullness: 27 ± 4 vs. 30 ± 4 mm) and were not influenced by time of day (a.m. vs. p.m. sessions). Men reported greater fasting hunger ($P = 0.04$) and desire to eat ($P = 0.03$) than women. Fullness ratings were not influenced by sex.

Fasting hunger (ICC(3,1) = 0.584, CCC = 0.575) and desire to eat (ICC(3,1) = 0.496, CCC = 0.487) demonstrated fair reliability and fullness ratings had good reliability (ICC(3,1) = 0.630, CCC = 0.622).

Correlations

No significant linear correlations among fasting appetite ratings and fasting brain responses were observed. Fasting brain responses were not linearly correlated with body mass index (data not shown).

Discussion

The purpose of this study was to assess the reproducibility of fasting-state brain responses to visual food stimuli in select reward-associated brain regions. Visual food stimuli elicited significant responses in all *a priori* regions of interest except for the left putamen on Day 1 and the right caudate on Day 2. In confirmation of our hypothesis, mean brain responses in *a priori* regions of interest were not different on the two testing days, with the exception of the left insula and right amygdala. Conversely and contrary to our hypothesis, no regions demonstrated good or excellent test-retest reliability. The left orbitofrontal cortex response had fair reliability, but reliabilities for the other 9 *a priori* regions of interest were poor or unreliable. For comparison, appetite ratings demonstrated fair (hunger, desire to eat) to good (fullness) test-retest reliability. These results are consistent with previous work indicating some degree of within subject variability in appetite ratings under similar experimental conditions, especially when using ratings from a single time point (30,31). Even though reliabilities for appetite ratings were higher than for fMRI-measured brain responses, it is possible that within subject variability in perceptions of appetite contributed to the relatively poor reliability observed in brain responses.

While we are unaware of extant data regarding the test-retest reliability of brain reward responses to visual food stimuli, a number of studies have assessed the reliability of reward responses using alternative fMRI study designs (29,32,33). In a group of alcohol-dependent individuals, reliabilities for responses to visual alcohol cues were poor to fair in the left ventral and dorsal striatum, but excellent in the right ventral and dorsal striatum (32). Reported ICCs for anticipatory reward responses were poor to fair in two event-related study designs utilizing monetary reward paradigms (29,33). Furthermore, a review of fMRI reliability studies investigating a range of study designs, outcomes, and brain regions found that the mean ICC across fMRI studies is 0.50 and thus demonstrate fair reliability overall (34). This suggests that the relatively poor reliability observed in the current study is not isolated to this single study or even to reward responses, but is rather more broadly characteristic of fMRI research.

On the other hand, the current study and previous studies (29,32,33) have reported that mean or group-level reward responses are relatively consistent. This general lack of good to excellent test-retest reliability ratings but good group-level consistency in fMRI-measured reward responses has important implications for the design and interpretation of fMRI-based intervention studies. Consistent group-level results suggest that parallel group design studies – without repeated fMRI scanning of subjects – are appropriate for fMRI-based studies of reward processing. Longer-term interventions that induce a physiological adaptation (e.g. weight loss, exercise training) may also cause group shifts in reward responses from pre- to post-intervention testing that could be detected by fMRI. Crossover intervention studies may be appropriate given that certain criteria are considered. For example, it is critical that investigators include chronological testing order as a factor in their statistical models and utilize standard randomization procedures to ensure that potential intervention effects cannot be explained by day-to-day within-subject variability in responses.

The test-retest reliability of a measurement also influences the maximum observable correlations (ρ_o) among study outcomes. The ρ_o between two variables a and b is theoretically limited by their reliabilities (R_a and R_b) such that the ρ_o is equal to the product of the true correlation (ρ_t) and the square root of the product of R_a and R_b .

$$\rho_o = \rho_t \sqrt{R_a R_b} \quad (13)$$

Using this equation, the likelihood of observing significant correlations with measurements demonstrating poor reliability or complete unreliability is severely compromised. Until the reliability of fMRI measurements can be improved, there is limited translational potential for developing inexpensive, effective tools that correlate well with fMRI-measured reward responses at the individual clinical patient level.

A primary strength of this study is the large sample size ($n = 28$) relative to previously published fMRI test-retest reliability assessments, which commonly have sample sizes of < 10 subjects (34). The present study was conducted in a group individuals with overweight and obesity, which is an appropriate group for nutrition-based interventions. Finally, factors

likely to influence fMRI signal detection and reliability were controlled by strictly adhering to experimental procedures and utilizing appropriate statistical controls.

There are a number of limitations in study design that could have influenced the test-retest reliability results. For example, the length of time between fMRI assessments ranged from 3 to 35 days, which may have influenced reliability ratings. Subjects were asked not to engage in moderate to vigorous exercise or consume alcohol for at least 48 hours, but these behaviors were not documented. Subjects who completed p.m. assessments were provided with a standard lunch on testing days, but dinner was not controlled the evening prior to scanning for a.m. sessions. Lastly, we used fasting-state fMRI and appetite assessments from 3 randomized controlled trials for this analysis. Postprandial fMRI scans were completed as part of those trials, and it is possible that subjects could have become habituated to the visual stimuli. However, the relatively consistent and significant mean brain responses on both testing days in most regions of interest suggests that habituation effects were minimal on average. Habituation effects were not directly assessed in the current study, which raises the possibility that individual differences in habituation among study participants impacted reliability ratings. While we are cognizant of the limitations of our approach, the results of current study raise an important and understudied aspect fMRI-based research, which is becoming increasingly prevalent in the study of ingestive behavior.

The use of ICC for test-retest reliability assessments also has some limitations, notably that a wider range of measurements/responses may result in a higher ICC and that non-normally distributed data may influence the resulting rating (35). Supplemental Figures S1-S10 demonstrate the degree of individual variability in brain responses from Day 1 to Day 2, which supports the conclusion of relatively poor reliability. Also, the CCC index produces robust results with non-normal distributions (27). Similar ICC and CCC ratings observed in the current study lend greater confidence to the results.

In conclusion, fMRI-measured brain responses to visual food stimuli in this group of adults with overweight or obesity demonstrated relatively consistent mean results but considerable within-subject variability on 2 days with similar experimental conditions. These findings have important implications relating to experimental design, sample size determination, and observable correlations among study outcomes. Future studies with more rigorous controls with regard to dietary intake, physical activity, and length of time between fMRI scanning should be conducted to determine to what extent these factors influence the test-retest reliability of fMRI-measured brain responses to visual food stimuli.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank the research subjects for their participation and dedication to the study. We are also grateful to Lexie Staten for her assistance as a secondary MRI operator and Amy Wright, RD for designing and preparing study meals.

Funding: NIH Indiana Clinical and Translational Sciences Institute, Clinical Research Center (Grant # UL1TR001108); USDA 2011-38420-20038; and Egg Nutrition Center/American Egg Board. Financial supporters had no role in the design and conduct of the study or collection, analysis, and interpretation of the data.

References

- Berthoud HR, Lenard NR, Shin AC. Food reward, hyperphagia, and obesity. *Am J Physiol Regul Integr Comp Physiol*. 2011; 300:R1266–77. [PubMed: 21411768]
- Tang DW, Fellows LK, Small DM, Dagher A. Food and drug cues activate similar brain regions: a meta-analysis of functional MRI studies. *Physiol Behav*. 2012; 106:317–24. [PubMed: 22450260]
- Burger KS, Berner LA. A functional neuroimaging review of obesity, appetitive hormones and ingestive behavior. *Physiol Behav*. 2014; 136:121–7. [PubMed: 24769220]
- Pursey KM, Stanwell P, Callister RJ, Brain K, Collins CE, Burrows TL. Neural responses to visual food cues according to weight status: a systematic review of functional magnetic resonance imaging studies. *Front Nutr*. 2014; 1:7. [PubMed: 25988110]
- Cornier MA, Melanson EL, Salzberg AK, Bechtell JL, Tregellas JR. The effects of exercise on the neuronal response to food cues. *Physiol Behav*. 2012; 105:1028–34. [PubMed: 22155218]
- Crabtree DR, Chambers ES, Hardwick RM, Blannin AK. The effects of high-intensity exercise on neural responses to images of food. *Am J Clin Nutr*. 2014; 99:258–67. [PubMed: 24305681]
- Evero N, Hackett LC, Clark RD, Phelan S, Hagobian TA. Aerobic exercise reduces neuronal responses in food reward brain regions. *J Appl Physiol*. 2012; 112:1612–9. [PubMed: 22383502]
- Leidy HJ, Lepping RJ, Savage CR, Harris CT. Neural responses to visual food stimuli after a normal vs. higher protein breakfast in breakfast-skipping teens: a pilot fMRI study. *Obes Silver Spring*. 2011; 19:2019–25.
- Leidy HJ, Ortinau LC, Douglas SM, Hoertel HA. Beneficial effects of a higher-protein breakfast on the appetitive, hormonal, and neural signals controlling energy intake regulation in overweight/obese, “breakfast-skipping,” late-adolescent girls. *Am J Clin Nutr*. 2013; 97:677–88. [PubMed: 23446906]
- Baicy K, London ED, Monterosso J, Wong ML, Delibasi T, Sharma A, Licinio J. Leptin replacement alters brain response to food cues in genetically leptin-deficient adults. *Proc Natl Acad Sci U A*. 2007; 104:18276–9.
- Malik S, McGlone F, Bedrossian D, Dagher A. Ghrelin modulates brain activity in areas that control appetitive behavior. *Cell Metab*. 2008; 7:400–9. [PubMed: 18460331]
- Guthoff M, Grichisch Y, Canova C, Tschritter O, Veit R, Hallschmid M, Haring HU, Preissl H, Hennige AM, Fritsche A. Insulin modulates food-related activity in the central nervous system. *J Clin Endocrinol Metab*. 2010; 95:748–55. [PubMed: 19996309]
- Fleiss, JL. *The Design and Analysis of Clinical Experiments*. John Wiley & Sons, Inc; New York, NY: 1986.
- Institute of Medicine (U.S.). Panel on Macronutrients. Institute of Medicine (U.S.). Standing Committee on the Scientific Evaluation of Dietary Reference Intakes. Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein, and amino acids. Vol. xxv. National Academies Press; Washington, DC: 2005. p. 1331
- Stubbs RJ, Hughes DA, Johnstone AM, Rowley E, Reid C, Elia M, Stratton R, Delargy H, King N, Blundell JE. The use of visual analogue scales to assess motivation to eat in human subjects: a review of their reliability and validity with an evaluation of new hand-held computerized systems for temporal tracking of appetite ratings. *Br J Nutr*. 2000; 84:405–15. [PubMed: 11103211]
- Marsh-Richard DM, Hatzis ES, Mathias CW, Venditti N, Dougherty DM. Adaptive Visual Analog Scales (AVAS): a modifiable software program for the creation, administration, and scoring of visual analog scales. *Behav Res Methods*. 2009; 41:99–106. [PubMed: 19182128]
- Burger KS, Cornier MA, Ingebrigtsen J, Johnson SL. Assessing food appeal and desire to eat: the effects of portion size & energy density. *Int J Behav Nutr Phys Act*. 2011; 8:101. [PubMed: 21943082]
- Cornier MA, Von Kaenel SS, Bessesen DH, Tregellas JR. Effects of overfeeding on the neuronal response to visual food cues. *Am J Clin Nutr*. 2007; 86:965–71. [PubMed: 17921372]

19. Cornier MA. The effects of overfeeding and propensity to weight gain on the neuronal responses to visual food cues. *Physiol Behav.* 2009; 97:525–30. [PubMed: 19328211]
20. Cornier MA, Salzberg AK, Endly DC, Bessesen DH, Rojas DC, Tregellas JR. The effects of overfeeding on the neuronal response to visual food cues in thin and reduced-obese individuals. *PLoS One.* 2009; 4:e6310. [PubMed: 19636426]
21. Cornier MA, Salzberg AK, Endly DC, Bessesen DH, Tregellas JR. Sex-based differences in the behavioral and neuronal responses to food. *Physiol Behav.* 2010; 99:538–43. [PubMed: 20096712]
22. Cornier MA, McFadden KL, Thomas EA, Bechtell JL, Eichman LS, Bessesen DH, Tregellas JR. Differences in the neuronal response to food in obesity-resistant as compared to obesity-prone individuals. *Physiol Behav.* 2013; 110:111–122.
23. Cornier MA, McFadden KL, Thomas EA, Bechtell JL, Bessesen DH, Tregellas JR. Propensity to obesity impacts the neuronal response to energy imbalance. *Front Behav Neurosci.* 2015; 9:52. [PubMed: 25767441]
24. Peirce JW. Generating Stimuli for Neuroscience Using PsychoPy. *Front Neuroinform.* 2008; 2:10. [PubMed: 19198666]
25. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res.* 1996; 29:162–73. [PubMed: 8812068]
26. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979; 86:420–8. [PubMed: 18839484]
27. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics.* 1989; 45:255–68. [PubMed: 2720055]
28. Cicchetti DV. The precision of reliability and validity estimates re-visited: distinguishing between clinical and statistical significance of sample size requirements. *J Clin Exp Neuropsychol.* 2001; 23:695–700. [PubMed: 11778646]
29. Plichta MM, Schwarz AJ, Grimm O, Morgen K, Mier D, Haddad L, Gerdes AB, Sauer C, Tost H, Esslinger C, et al. Test-retest reliability of evoked BOLD signals from a cognitive-emotive fMRI test battery. *Neuroimage.* 2012; 60:1746–58. [PubMed: 22330316]
30. Raben A, Tagliabue A, Astrup A. The reproducibility of subjective appetite scores. *Br J Nutr.* 1995; 73:517–30. [PubMed: 7794869]
31. Flint A, Raben A, Blundell JE, Astrup A. Reproducibility, power and validity of visual analogue scales in assessment of appetite sensations in single test meal studies. *Int J Obes Relat Metab Disord J Int Assoc Study Obes.* 2000; 24:38–48.
32. Schacht JP, Anton RF, Randall PK, Li X, Henderson S, Myrick H. Stability of fMRI striatal response to alcohol cues: a hierarchical linear modeling approach. *Neuroimage.* 2011; 56:61–8. [PubMed: 21316465]
33. Fliessbach K, Rohe T, Linder NS, Trautner P, Elger CE, Weber B. Retest reliability of reward-related BOLD signals. *Neuroimage.* 2010; 50:1168–76. [PubMed: 20083206]
34. Bennett CM, Miller MB. How reliable are the results from functional magnetic resonance imaging? *Ann N Acad Sci.* 2010; 1191:133–55.
35. Muller R, Buttner P. A critical discussion of intraclass correlation coefficients. *Stat Med.* 1994; 13:2465–76. [PubMed: 7701147]

Study Importance

- The brain's reward system influences ingestive behavior and subsequently obesity risk.
- Previous data indicate that visual food cues produce robust responses in reward-associated brain regions that are greater in adults with obesity compared to those with a normal body weight.
- In the current study, brain responses to visual food stimuli measured by functional magnetic resonance imaging show considerable within-subject variability in adults with overweight or obesity that needs to be considered when designing future studies.

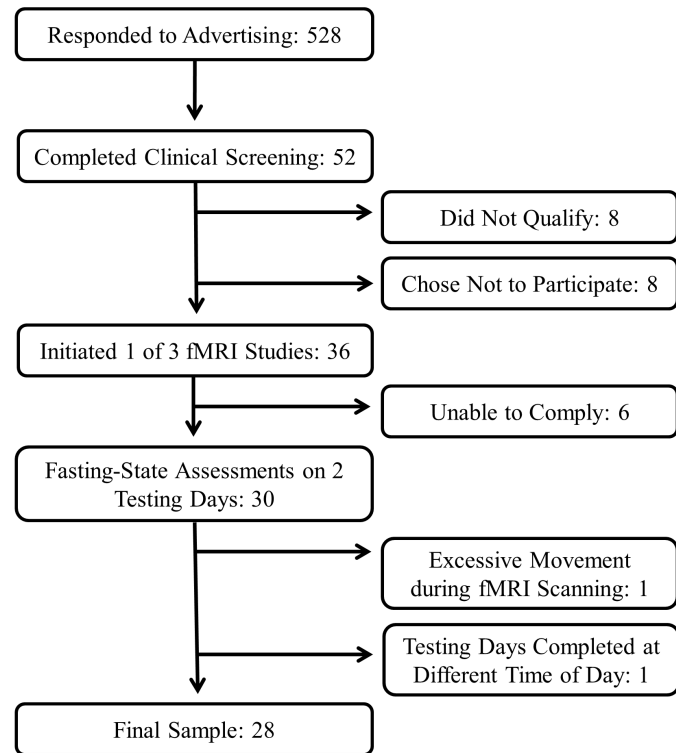


Figure 1.
Study recruitment flow diagram

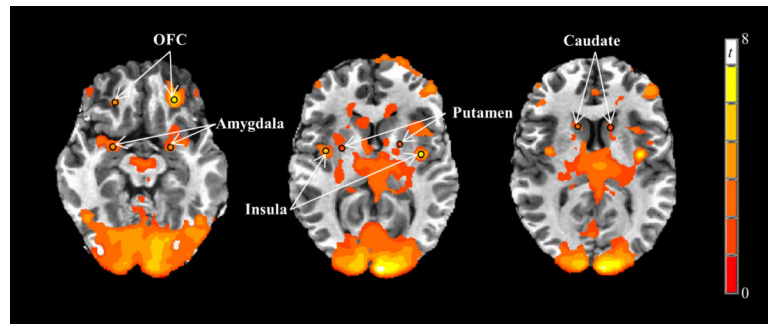


Figure 2.

Fasting-state brain responses to visual food stimuli on Day 1. Greater responses to visual food stimuli vs. nonfood stimuli (PROC TTEST, SAS, Version 9.3; $P < 0.005$) were observed in 9 of 10 *a priori* regions of interest (left putamen response was not significant). Black circles represent functional regions of interest with 3mm radii within *a priori* brain regions of interest with known reward functions. Images are in the axial plane and left side of the figure corresponds to the right side of the body and vice versa. Display threshold: $p < 0.001$ (uncorrected), minimum cluster size of 250 voxels.

Abbreviations: OFC, orbitofrontal cortex

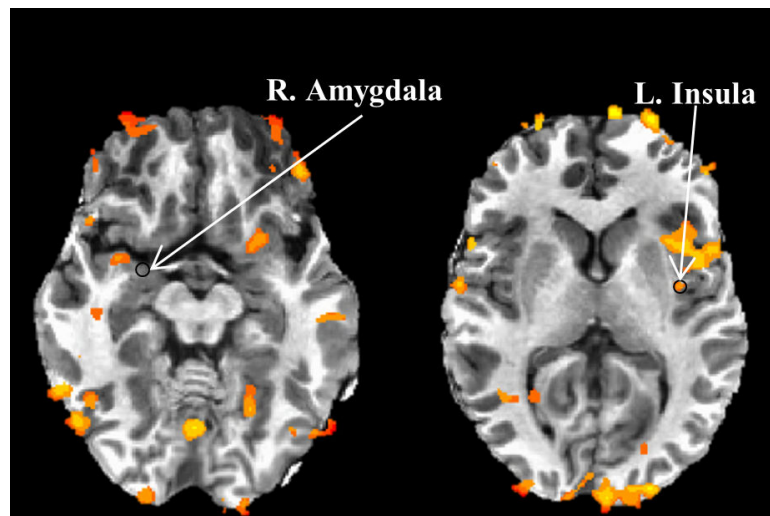


Figure 3.

Voxel-wise subtraction of Day 1 vs. Day 2 fasting-state brain responses. The ANOVA model (PROC MIXED, SAS, Version 9.3) indicated greater responses to visual food stimuli in the right amygdala on Day 1 vs. Day 2 and a trend for greater Day 1 left insula responses. The voxel-wise subtraction also suggests an attenuation of left insula responses on Day 2, but does not support a difference in Day 1 vs. Day 2 right amygdala responses. Images are in the axial plane and left side of the figure corresponds to the right side of the body and vice versa. Display threshold: $p < 0.05$ (uncorrected), minimum cluster size of 250 voxels.

Table 1Subject characteristics^{*I*}

Parameter	n = 28
Male/Female	12/16
Age (y)	27 ± 1
Body Mass (kg)	86.3 ± 2.4
Body Mass Index (kg/m ²)	29.4 ± 0.8

^{*I*}Values are mean ± SEM.

Abbreviations: fMRI, functional magnetic resonance imaging

Table 2

Responses to visual food cues compared to nonfood cues in the fasting-state on Day 1 and Day 2

Brain Region	MNI coordinates ¹			Day 1		Day 2	
	x	y	z	T value ²	P value ³	T value	P value
Insula (L)	-38	-7	6	10.83	< 0.0001	6.34	< 0.0001
Insula (R)	39	-4	4	6.39	< 0.0001	8.07	< 0.0001
Amygdala (L)	-23	0	-17	4.71	< 0.0001	4.67	< 0.0001
Amygdala (R)	24	0	-18	5.54	< 0.0001	3.32	0.0025
OFC (L)	-25	35	-18	6.15	< 0.0001	6.66	< 0.0001
OFC (R)	23	33	-20	4.74	< 0.0001	7.47	< 0.0001
Caudate (L)	-13	15	4	3.38	0.0022	3.58	0.0013
Caudate (R)	14	14	6	3.66	0.0011	2.90	0.0071
Putamen (L)	-22	8	2	2.84	0.0086	3.67	0.0010
Putamen (R)	27	5	0	4.20	0.0003	3.52	0.0015

¹ Stereotactic coordinates in MNI space for local maxima for the food vs. nonfood contrast on Day 1 within each *a priori* brain region of interest.

² T values reported for results of single sample *t*-tests for comparison of mean β coefficient of all voxels in each region for food vs. nonfood contrast to zero.

³ Uncorrected *P* values.

Single sample Student's *t*-tests (PROC TTEST, SAS, Version 9.3) indicated that visual food stimuli elicited greater brain responses compared to nonfood stimuli in all *a priori* regions of interest except the left putamen on Day 1 and the right caudate on Day 2. A Bonferroni correction was applied for determination of statistical significance ($\alpha = 0.05 / 10 \text{ ROI} = 0.005$).

Abbreviations: MNI, Montreal Neurological Institute; OFC, orbitofrontal cortex; L, left; R, right.

Table 3Comparison of fasting-state responses to visual food cues on Day 1 and Day 2¹

Brain Region	Day 1 Response (β)	Day 2 Response (β)	P value ²
Insula (L)	$1.20 \times 10^{-3} \pm 1.11 \times 10^{-4}$	$9.37 \times 10^{-4} \pm 1.50 \times 10^{-4}$	0.076
Insula (R)	$9.74 \times 10^{-4} \pm 1.52 \times 10^{-4}$	$9.19 \times 10^{-4} \pm 1.15 \times 10^{-4}$	0.74
Amygdala (L)	$1.34 \times 10^{-3} \pm 2.85 \times 10^{-4}$	$1.19 \times 10^{-3} \pm 2.55 \times 10^{-4}$	0.62
Amygdala (R)	$1.28 \times 10^{-3} \pm 2.30 \times 10^{-4}$	$5.78 \times 10^{-4} \pm 1.83 \times 10^{-4}$	0.014 [*]
OFC (L)	$1.77 \times 10^{-3} \pm 2.87 \times 10^{-4}$	$1.59 \times 10^{-3} \pm 2.35 \times 10^{-4}$	0.46
OFC (R)	$1.36 \times 10^{-3} \pm 2.87 \times 10^{-4}$	$1.29 \times 10^{-3} \pm 1.78 \times 10^{-4}$	0.81
Caudate (L)	$5.17 \times 10^{-4} \pm 1.53 \times 10^{-4}$	$4.38 \times 10^{-4} \pm 1.20 \times 10^{-4}$	0.69
Caudate (R)	$5.52 \times 10^{-4} \pm 1.51 \times 10^{-4}$	$4.58 \times 10^{-4} \pm 1.32 \times 10^{-4}$	0.66
Putamen (L)	$3.66 \times 10^{-4} \pm 1.19 \times 10^{-4}$	$3.51 \times 10^{-4} \pm 9.58 \times 10^{-5}$	0.97
Putamen (R)	$5.82 \times 10^{-4} \pm 1.39 \times 10^{-4}$	$4.91 \times 10^{-4} \pm 1.34 \times 10^{-4}$	0.69

¹Values are reported as mean β coefficient of all voxels in each region for food vs nonfood contrast. Variability is reported as \pm SEM.

²P values for comparison of Day 1 and Day 2 responses.

^{*}Indicates greater responses to visual food stimuli on Day 1 compared Day 2. Repeated measures ANOVA (PROC MIXED, SAS, Version 9.3) indicate that visual food stimuli elicited greater responses right amygdala on Day 1 compared to Day 2. A non-significant trend for greater responses on Day 1 was observed in the left insula.

Abbreviations: OFC, orbitofrontal cortex; L, left; R, right.

Table 4

Reliabilities of fasting state responses to visual food cues

Brain Region	ICC(3,1)	CCC	Rating ^I
Insula (L)	0.389	0.382	Poor
Insula (R)	0.245	0.239	Poor
Amygdala (L)	0.390	0.381	Poor
Amygdala (R)	0.153	0.148	Poor
OFC (L)	0.575	0.566	Fair
OFC (R)	0.306	0.297	Poor
Caudate (L)	−0.005	−0.004	Unreliable
Caudate (R)	−0.150	−0.145	Unreliable
Putamen (L)	0.102	0.099	Poor
Putamen (R)	−0.405	−0.386	Unreliable

^IReliability ratings determined from ICC(3,1) and CCC analyses were interpreted as: < 0.40: poor reliability, 0.40 – 0.59: fair reliability, 0.60 – 0.74: good reliability, 0.75: excellent reliability. Negative ICCs were interpreted as being equivalent to zero and to represent complete unreliability.

ICC(3,1) and CCC were calculated using IBM SPSS Statistics, Version 22. Fasting state responses to visual food cues were mostly poor or completely unreliable, with only the left orbitofrontal cortex response demonstrating fair reliability.

Abbreviations: ICC(3,1), 2-way mixed models type intra-class correlation coefficient; CCC, Lin's concordance correlation coefficient; OFC, orbitofrontal cortex; L, left; R, right.